CEWES MSRC/PET TR/99-08

1998 CEWES MSRC PET Training Activities

by

Wayne Mastin

04h00699

Work funded by the DoD High Performance Computing Modernization Program CEWES Major Shared Resource Center through

Programming Environment and Training (PET)

Supported by Contract Number: DAHC94-96-C0002
Nichols Research Corporation

1998 CEWES MSRC PET Training Activities

Wayne Mastin*

March 4, 1999

1 Introduction

This report summarizes the training activities conducted through the PET program at the CEWES MSRC for the 1998 calendar year. These activities were organized by the PET on-site staff in collaboration with Syracuse University, PET lead university for training, and the members of the PET university team providing training classes to the CEWES MSRC user community.

The PET training program continued its evolution in 1998 with more emphasis on distance training technology and service to remote users. The first full-blown distance training class was offered through the Tango Interactive distance consulting system. This year has also seen continuous development of Tango as a distance education tool through its application as a vehicle for offering graduate computer science courses from Syracuse University to Jackson State University.

The PET training team supports not only training activities, but also provides logistic and technical support to all areas of the PET program. Examples include support for the Grid Generation Workshop in February and the CSM Workshop in November.

2 Training Classes

PET training is designed to assist the CEWES MSRC user in transitioning to new programming environments and efficiently using the present and future SPP (Scalable Parallel Processing) hardware acquired under the HPCM program. The training curriculum is a living document with new topics being added continually to keep up with the fast pace of research and development in the field of HPC. The curriculum contains courses in the following general categories.

- Parallel programming
- Architecture and software specific topics
- Visualization and performance
- CTA targeted classes, workshops, and forums

Table 1 gives a list of all classes taught during 1998 with the organization offering the class, the number of students attending the class, and the overall evaluation score of the class on a scale of 1 (poor) to 5 (excellent). Unless otherwise noted, the classes were held in the CEWES MSRC Training and Education Facility (TEF). From this table, we have the following totals.

*PET On-Site Team Lead, Nichols Research and Professor Emeritus, Mississippi State University

- Number of classes: 24
- Number of remote classes: 9
- Number of students: 250

Date	Class	Provider	Students	Evaluation
Jan. 15-16	Code Parallelization	Texas	19	4.0
Jan. 27-28	Performance Tools	Tennessee	8	4.0
Feb. 11	Code Optimization	Tennessee	11	4.3
Mar. 3-6	Parallel Programming Workshop[2]	CEWES MSRC	8	5.0
Mar. 16-20	Parallel Programming Workshop	CEWES MSRC	3	4.5
Apr. 29-May 1	MPI[1]	MSU	5	5.0
May 5-6	Paradyn/DYNA3D	LLNL	10	N/A
May 12-13	Grid Generation	MSU	10	3.8
May 19-21	Java	Syracuse	10	4.5
Jun. 1	Grid Generation[3]	MSU	10	N/A
Jun. 1	MPI[3]	MSU	22	N/A
Jun. 1	HPF[3]	Rice	12	N/A
Jun. 1	Performance Optimization[3]	Tennessee	20	N/A
Jun. 1	Finite Elements[3]	Texas	15	N/A
Jun 17-19	MPI	MSU	3	5.0
Jun. 22-26	Parallel Programming Workshop	CEWES MSRC	3	4.7
Aug. 3-4	Designing and Building Par. Prog.[1]	Rice	3	5.0
Aug. 6-7	Designing and Building Par. Prog.	Rice	6	4.0
Sep. 28	Fortran 90	OSC	21	4.2
Oct. 21	OpenMP	SGI/Cray	13	4.8
Oct. 28-30	PATRAN	MSC	8	4.5
Nov. 9-13	Parallel Programming Workshop	CEWES MSRC	4	4.8
Nov. 17-18	Tango	Syracuse	5	4.8
Dec. 7-11	Parallel Programming Workshop[4]	CEWES MSRC	12	N/A
Dec. 8-9	TotalView/Vampir	Tennessee	9	4.2

[1]Taught at NRL, Washington, DC
[2]Taught at AEDC, Arnold AFB, TN
[3]Taught at DoD HPC Users Group Conference, Houston, TX
[4]Taught at Groton, CT

Table 1: Classes taught from January through December 1998

The Fortran 90 class offered on September 28 was our first Tango-based class offered to remote users. The class was broadcast over the Internet to OSC, the class provider, and users at the ARL MSRC. Additional Tango classes have been scheduled for 1999.

Many of the training classes were recorded on VHS tapes and broadcast over Mbone. Information on availability of tapes may be obtained from CEWES MSRC HPC User Support by email at info-hpc@wes.army.mil or phone at 601-634-4400 (option 1) or 1-800-500-HPCC. Available tapes may also be ordered from the CEWES MSRC web site at
`http://www.wes.hpc.mil/msrc/training/registration/video_form.html`.

Descriptions for all training classes offered in 1998 appear in Appendix A. Information on classes taught in previous years can be found on the web at `http://www.wes.hpc.mil/msrc/training/past/`.

3 DoD HPC Users Group Conference

The CEWES MSRC PET program sponsored training activities at the DoD HPC Users Group Conference in Houston. Five training classes were held and are included in the list of classes in Table 1. These training classes had the largest attendance in the history of the User Group Conferences.

The PET program also sponsored a PET Training Colloquium on Distance Learning and Collaboration. The colloquium was organized by Prof. Geoffrey Fox, PET Academic Lead for Training and Collaboration/Communication, from Syracuse University. The speakers were Dr. Anoop Gupta of Microsoft Research, Dr. Don Johnson of the DoD Advanced Distributed Learning Initiative, and Prof. Fox. The moderator was Dr. Louis Turcotte, CEWES MSRC.

4 Seminars

The PET program offers seminars on an irregular basis. These are presentations by experts in their field and are designed to introduce the CEWES MSRC users to current research topics in HPC. A list of seminar presentations for 1998 appears in Appendix B.

5 HBCU/MI Efforts

Jackson State University and Syracuse University have worked together to refine the Tango Interactive technology for distance education. During the spring 1998 semester, Syracuse taught the JSU course CSC 499 Programming for the WEB. This was an undergraduate computer science programming course that covered mainly tools for creating and working with web pages. Topics covered included HTML, Java, CGI, and Perl. For the fall 1998 semester, Syracuse taught the JSU course CSC 539 Computational Science for Simulation Applications. This was a graduate computer science course. It covered numerical methods in scientific computing and included topics in scalable parallel computing, such as message passing. CSC 539 was made available in the TEF and the course was audited by three CEWES MSRC users.

Both of the courses were broadcast over the Internet from Syracuse to JSU using the DREN and the T1 line between CEWES MSRC and JSU. Course descriptions are included in Appendix C.

6 Conclusions

Since its inception, the CEWES MSRC PET training program has faced two challenges. One is to provide training in an anytime, anypace, anyplace environment. That goal has not been reached, but the PET program has continued to support efforts in remote training and distance education. Those efforts are now bearing fruit as can be seen from this report. The second challenge is to meet the needs of users faced with a rapid change in available hardware and software systems. The attempts to meet this challenge are evident by comparing the courses listed in Appendix A with the same list that appeared in the report of 1997 training activities [1]. Training is now offered on products like OpenMP that did not exist in 1997, and classes on topics such as C++ have been replaced by Java classes.

The scheduling of training classes and other events is coordinated with the MSRCs at ARL, ASC, and NAVO. Jointly sponsored activities, such as the Tango-based Fortran 90 class, will continue.

The training curriculum is driven by user needs and continues to evolve. The current schedule of training classes is on the web at `http://www.wes.hpc.mil/msrc/training/f_cewes.html`. Suggestions for future training are always welcome and can be made by contacting Dr. Wayne Mastin by email at mastin@nrcpet1.wes.hpc.mil or by phone at 601-634-3063.

References

[1] Wayne Mastin, "1997 CEWES MSRC PET Training Activities," CEWES MSRC PET TR 98-01, Vicksburg, MS, 1998.

A Training Classes

Techniques in Code Parllelization

Description

The techniques needed to parallelize a code are described. These includes partitioning, load balancing, preprocessing, and postprocessing. Examples of parallelization efforts carried out at the University of Texas will be given.

Method of Delivery

Lecture

Prerequisites

Basic knowledge of numerical methods and parallel processing

Organization Providing the Training

University of Texas

Target Audience

Users and developers of parallel codes for SPP systems

Class Duration

2 days

Workshop on Portable Parallel Performance Tools

Description

This workshop will cover the basics of tool-assisted performance analysis and tuning, as well as introduce a number of tools, both research and commercial, that are available on multiple parallel platforms. Both post-mortem analysis of trace files generated during program execution, and run-time analysis using dynamic instrumentation, will be covered. Tools to be covered include AIMS, MPE logging and nupshot, Pablo, Paradyn, and VAMPIR.

Method of Delivery

Lecture and lab

Prerequisites

Basic knowledge of parallel processing and familiarity with MPI

Organization Providing the Training

University of Tennessee

Target Audience

Engineers and scientists interested in learning to use portable performance tools for analyzing and tuning the performance of parallel applications using the Message Passing Interface (MPI).

Class Duration

2 days

Code Optimization

Description

The course will begin with a quick overview of the basics of performance and processor architecture. Then we will cover a wide variety of optimizations geared towards enhancing single processor performance. Topics will include efficient use of the memory hierarchy, functional units, amortizing loop overhead and dependency analysis. Common bottlenecks and caveats will be discussed as well as proposed solutions, and the logic behind them. After covering the single processor case, we will progress towards specific optimizations geared towards MPPs. Topics will include betters ways of data layout, appropriate granularity of computation and reducing communication and contention for resources. Lastly we will cover some details of each of the architectures in the SP, the T3E and the Origin 2000. Specific limitations of the respective architectures will be discussed as well as how they affect the various optimization techniques. In addition, each platform is shipped with programming tools designed to aid in the optimization process. The course will conclude with a quick survey of these tools and their usefulness to the application engineer.

Method of Delivery

Lecture and lab

Prerequisites

Some experience in high performance computing

Organization Providing the Training

University of Tennessee

Target Audience

Users and developers of parallel codes for MPP systems

Class Duration

3 days

Parallel Programming Workshop for Fortran Programmers

Description

The workshop will begin with a one-day lecture on strategy, tools, and examples in parallel programming. On the remaining days participants will work with their own codes. There will be no attempt to prescribe a particular solution to the problem of porting a code from the C90 to the scalable systems. Rather, the instructors will work with the user to find the best overall strategy, whether that best strategy is message passing via MPI or PVM, or data parallel via HPF or OpenMP. It may not be possible to parallelize a full blown application program in a week, but the process can get started and a continuing relationship can be established between the users and the parallelization experts at the CEWES MSRC.

Method of Delivery

Hands-on laboratory

Prerequisites

Basic knowledge of parallel processing and familiarity with the Fortran programming language. Participants must bring a candidate code to be parallelized. Familiarity with that code is helpful. The ideal code is one which is similar to one known to be parallelized, is important enough to be worth spending time on it, and yet is simple enough so it can be understood and the edit-compile-debug cycle isn't excessively long.

Organization Providing the Training

CEWES MSRC

Target Audience

Application programmmers who have codes they are interested in porting from the Cray C90 to the new scalable systems: the Cray T3E, the IBM SP, and the SGI Origin 2000.

Class Duration

3-5 days

Using the Message Passing Interface (MPI) Standard

Description

Message-Passing Interface (MPI) is the de facto standard for message-passing developed by the Message-Passing Interface Forum (MPIF). MPI provides many features needed to build portable, efficient, scalable, and heterogeneous message-passing code. These features include point-to-point and collective communication, support for datatypes, virtual topologies, process-group and communication context management, and language bindings for the FORTRAN and C languages. In this tutorial we will cover the important features supported by MPI with examples and illustrations. Also an introduction to extensions of MPI (MPI-2) and message-passing in real-time (MPI/RT) will also be provided.

Method of Delivery

Lecture and lab

Prerequisites

Basic knowledge of parallel processing and familiarity with C or FORTRAN programming language.

Organization Providing the Training

Mississippi State University

Target Audience

Developers of parallel libraries and parallel applications, whose aim is to develop programs/applications that provide portability and performance over a wide range of high performance computing systems (homogeneous as well as heterogeneous). Programmers who wish to improve their parallel programming skills will also find this tutorial useful.

Class Duration

3 day

Large Deformation Computational Structural Mechanics Applications on High Performance Computers using ParaDyn/DYNA3D

Description

This course will begin with a DYNA3D lecture reviewing the features added to the program since 1993. Some of the recent features include techniques for switching materials from rigid to deformable and back, new material models and equations of state, recent developments in element technology, and new contact methods. This lecture will include time for questions and answers about modeling and using any of the features in DYNA3D. The MSRC will provide attendees a summary of steps required for submitting batch jobs to run parallel problems on the Origin 2000, Cray T3E and IBM SP. This will include the design of script files for the batch system, a discussion of the batch queues, and running the batch utilities to follow the progress of a job. The ParaDyn lecture will feature discussions on the automated software for domain decomposition, running the ParaDyn program, post-processing the results for visualization, and the performance on parallel computers. Techniques for efficiently handling contact boundary conditions and future parallel capability releases will be discussed. The lectures will finish with a discussion of applications illustrating the power of parallel computers in modeling problems of DoD interest. On the second day the instructor will demonstrate a sample problem preparation and execution of a ParaDyn calculation on one of the parallel systems at CEWES MSRC. Attendees will be able to run their own examples and work with the instructor directly at this time.

Method of Delivery

Lecture, demonstation and hands-on exercises

Prerequisites

Experience using various features of the DYNA3D code and experience as a structural analyst

Organization Providing the Training

Lawrence Livermore National Laboratory

Target Audience

Advanced-level DYNA3D users desiring to run large models (over 100,000 elements) on parallel computers

Class Duration

2 days

Grid Generation for Complex Configurations

Description

This course will cover an in-depth review of the current state-of-the-art and state-of-practice in geometry/grid generation applicable to complex problems. A step-by-step process starting from the initial CAD definition or drawing of a configuration and proceeding to the generation of a curvilinear, hexahedral or cartesian grid and grid adaptation techniques will be presented in detail. Demonstrations and hands-on computer lab exercises will be conducted to explore the use of GUM-B, VGRID, CAGI, GENIE++, TrueGrid, PMAG, CUBIT, and Hybrid2d systems for practical applications of interest to CEWES MSRC users.

Method of Delivery

Lecture, demonstation and hands-on exercises

Prerequisites

Basic knowledge in numerical methods

Organization Providing the Training

Mississippi State University

Target Audience

Novice users needing to construct computational grids for numerical simulation

Class Duration

2 days

Java for Scientific Computing

Description

The objective of this course is to provide the participant with a) an understanding of the high performance computing architecture, including the World Wide Web for visualization, b) an overview of the Java language and its capabilities, and c) enough programming details to do some examples.

Method of Delivery

Lecture and lab

Prerequisites

Programming experience in a language such as C, C++, or Fortran

Organization Providing the Training

Syracuse University

Target Audience

Scientists, computer scientists, and engineers

Class Duration

3 days

High Performance Fortran (HPF) in Practice

Description

This course will introduce programmers to the most important features of HPF, including features inherited from Fortran 90, the data parallel FORALL statement and INDEPENDENT assertion, and data mapping by ALIGN and DISTRIBUTE directives. The instructor will illustrate how these features can be used in practice on algorithms for scientific computation such as LU decomposition and the conjugate gradient method.

Method of Delivery

Lecture

Prerequisites

Some knowledge of Fortran 77 (or a similar imperative sequential programming language); a basic knowledge of scientific computation and/or parallelism is also useful, but not essential

Organization Providing the Training

Rice University

Target Audience

Researchers and practitioners who are interested in applying data-parallel computation to scientific programs

Class Duration

1 day

Performance Optimization

Description

This course will focus on the optimization of numeric intensive codes for HPC systems. The course will begin with a quick overview of the basics of performance and processor architecture. Then it will cover a wide variety of optimizations geared towards enhancing processor performance. Topics will include efficient use of the memory hierarchy, functional units, amortizing loop overhead and dependency analysis. Common bottlenecks and caveats will be discussed as well as proposed solutions, and the logic behind them.

Method of Delivery

Lecture and lab

Prerequisites

Some experience in high performance computing

Organization Providing the Training

University of Tennessee

Target Audience

Users and developers of codes for HPC systems

Class Duration

1 day

Topics in Finite Element Methodology for Nonlinear Problems

Description

This course is broadly structured to cover different types of applications from structures to fluids to heat transfer and coupled problems that are of general interest to DoD. Methodology rather than specific applications is stressed. Topics covered include algorithms, nonlinear solution strategies, and integrating solution with adaptive refinement.

Method of Delivery

Lecture

Prerequisites

A background in engineering or applied sciences and some experience in using finite element methods

Organization Providing the Training

University of Texas

Target Audience

Users needing to apply finite element methods to the solution of large scale nonlinear problems in computational engineering

Class Duration

1 day

A Tutorial on Designing and Building Parallel Programs

Description

In this tutorial, the instructors will provide a comprehensive introduction to the techniques and tools used to write parallel programs. First, the instructors will introduce principles of parallel program design, touching upon relevant topics in architecture, algorithms, and performance modeling. Examples from well-established parallel programming systems (HPF and MPI) will be included. After the basic material is covered, we will examine two new programming systems for parallel machines, OpenMP and PETSc.

Method of Delivery

Lecture and lab

Prerequisites

Fortran or C programming

Organization Providing the Training

Rice University

Target Audience

Scientists and engineers needing to write parallel programs in an efficient, portable way

Class Duration

2 days

An Introduction to the Fortran 90

Description

This course is aimed at introducing engineers and scientists familiar with Fortran 77 to the new features and capabilities available in Fortran 90. These new features include free form source code, the CASE control structure, the ability to create new data types, modules (similar to C++ classes), array processing shortcuts, dynamic memory allocation, pointers, improved I/O handling, and a host of new intrinsic functions. Source code compatibility between Fortran 77 and Fortran 90 will also be discussed.

Method of Delivery

Lecture

Prerequisites

Attendees should be familiar with Fortran 77 and numerical methods. Familiarity with other programming languages such as C or C++ may be helpful but is not required.

Organization Providing the Training

Ohio Supercomputer Center

Target Audience

Fortran programmers

Class Duration

1 day

Scalable OpenMP Programming on Origin2000

Description

This is an advanced course. Topics to be covered are:

- Overview of OpenMP programming model
- Review of execution model
- Moving beyond incremental parallelization mode
- Domain decomposition
- Comparisons with message passing
- Performance optimization on Origin2000
- Preview of OpenMP C/C++ specification

Method of Delivery

Lecture

Prerequisites

Some familiarity with OpenMP

Organization Providing the Training

SGI/Cray

Target Audience

Origin2000 programmers

Class Duration

1 day

Introduction to MSC/PATRAN - Modeling for Design Analysis

Description

This is an introductory course for new and/or infrequent MSC/PATRAN user. Students will master the basic skills required to use MSC/PATRAN in a typical MCAE application. This course emphasizes practical skills development through comprehensive, hands-on laboratory sessions. Students will learn to build analysis models using MSC/PATRAN, by defining material properties, creating boundary conditions, and submitting their problems for analysis and post-processing the results using a variety of graphical formats. Specific topics such as CAD integration, geometry editing, meshing, grouping, and customization will be covered. Users of all FEA codes are encouraged to attended since MSC/PATRAN supports all the popular FEA codes such as MSC/NASTRAN, MSC/DYTRAN, HKS/ABAQUS, ANSYS, LS-DYNA and many more.

Method of Delivery

Lecture and hands-on laboratory

Prerequisites

Knowledge of finite element analysis desirable, but not required.

Organization Providing the Training

MacNeal-Schwendler Corporation

Target Audience

PATRAN users

Class Duration

3 days

Tango for Remote Consulting

Description

NPAC's Tango is a Web collaboratory. The system extends capabilities of Web browsers towards a fully interactive, multimedia, collaborative environment. Tango is also a framework for building collaboratory systems. In this tutorial we will instruct how to use Tango and will cover applications of Tango for remote consulting, including all the critical software development phases: coding, compiling, testing and debugging, result analysis.

Method of Delivery

Lecture and lab

Prerequisites

Basic knowledge of C programming language and GNU debugger (gdb). Familiarity with Emacs or XEmacs is desired.

Organization Providing the Training

Syracuse University

Target Audience

Programmers who are members of geographically separated teams working together on complex short-term (consulting) or long-term (software development) efforts.

Class Duration

2 days

Parallel Debugging and Performance Analysis Tools: TotalView and Vampir

Description

The goal of this course is to introduce parallel application developers to parallel debugging and performance analysis tools available on CEWES MSRC platforms, and to provide more in-depth coverage of the TotalView debugger and Vampir performance analysis tool. The course will cover the basics of using the tools as well as provide pointers to further information. A lab session will include practice on using the tools on some example programs. Debuggers to be covered include Dolphin TotalView 3.8 for the SGI/Cray Origin 2000 and IBM SP, Cray TotalView for the Cray T3E, SGI dbx for the Origin 2000, and pdbx for the IBM SP. Dolphin TotalView has a graphical interface while dbx and pdbx provide command-line debugging interfaces. The Cray version of TotalView for the Cray T3E has both graphical and command-line interfaces. An overview will be given of the various performance analysis tools available on CEWES MSRC platforms, but the performance analysis portion of the course will focus in detail on the Vampir tool which has been recently acquired and is now available on CEWES MSRC machines.

Method of Delivery

Lecture and hands-on laboratory

Prerequisites

Some parallel programming experience

Organization Providing the Training

University of Tennesse

Target Audience

Parallel application developers who use Fortran 77/90, C, C++, HPF, MPI

Class Duration

2 days

B Seminars

Scalable Computing using the Bulk Synchronous Parallel Model

by

Dr. Jon Hill
Research Officer, Oxford Parallel
Oxford University
and
Prof. Bill McColl
Professor of Computing Science and Chairman of Oxford Parallel
Oxford University

9:00 am, Friday, March 20, 1998
P. K. Senter Conference Room
CEWES MSRC
and
3:00 pm, Friday, March 20, 1998
Peoples Science Building, Room 209
Jackson State University

Managing Scientific Data with HDF

by

Dr. Michael Folk
National Center for Supercomputing Applications (NCSA)
University of Illinois

10:00 am, Friday, July 10, 1998
P. K. Senter Conference Room
CEWES MSRC

Web-Based Instruction

by

Prof. Geoffrey Fox
Director, Northeast Parallel Architectures Center (NPAC)
Syracuse University

8:30 am, Wednesday, September 16, 1998
P. K. Senter Conference Room
CEWES MSRC

C Distance Education Courses

CSC 499 Programming for the WEB

Description

This programming course covers basic programming languages and skills to provide a basis for the further study of Web software applications. Students learn Web architecture and Web interfacing mechanisms through the Common Gateway Interface (CGI). The main part of the course concentrates on the use of the Java programming language for Web user interfaces and for distributed computing.

Method of Delivery

Tango from Syracuse to JSU

Prerequisites

An introductory course in computer programming

Organization Providing the Training

Syracuse University

Target Audience

Upper level undergraduate students in computer science

CSC 539 Computational Science for Simulation Applications

Description

This course is the graduate level introductory course in the discipline of Computational Science (the computer simulation of natural systems). This course is designed to teach the basic tools from mathematics and computer science that are needed to give computational solutions to scientific and engineering problems. Topics may include matrix methods, random numbers and Monte Carlo methods, numerical methods for ordinary and partial differential equations, and optimization techniques.

Method of Delivery

Tango from Syracuse to JSU

Prerequisites

Computer programming experience

Organization Providing the Course

Syracuse University

Target Audience

Advanced undergraduate computer science students